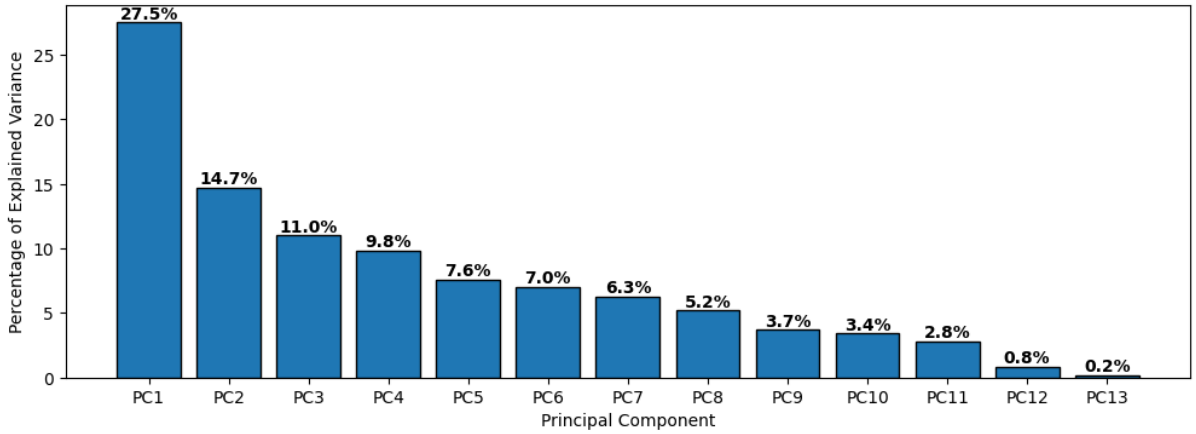
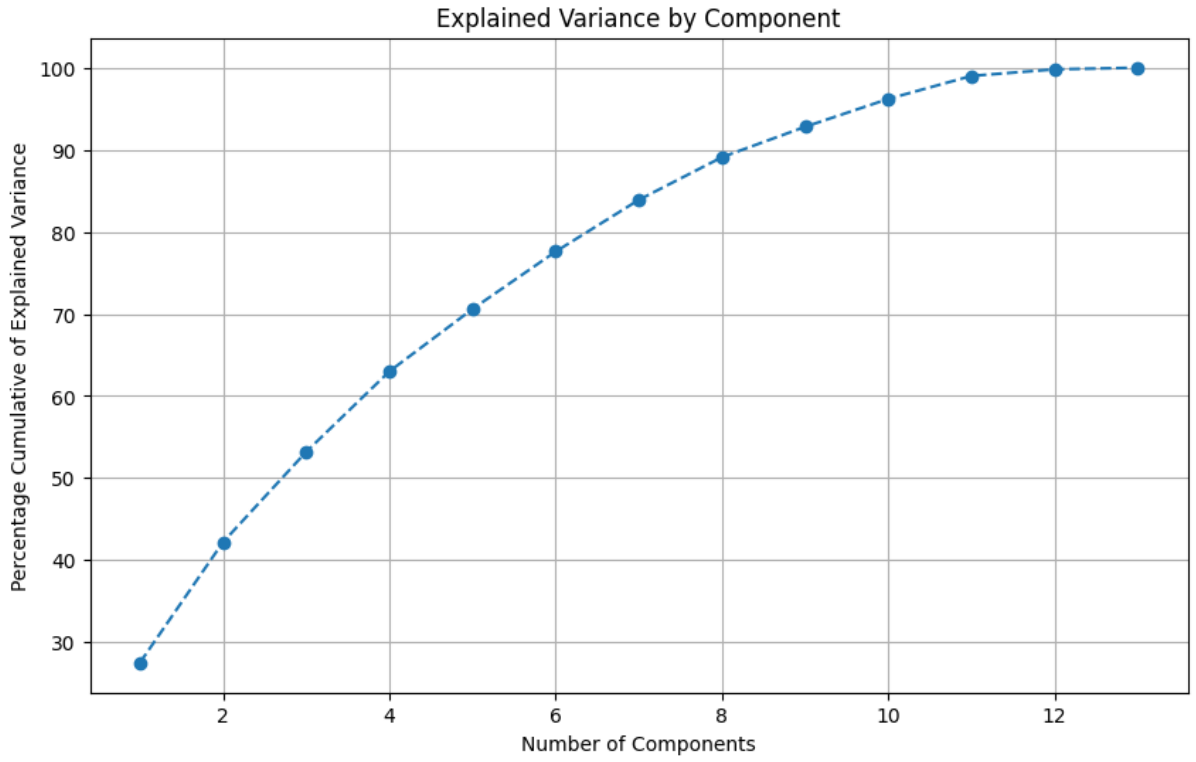
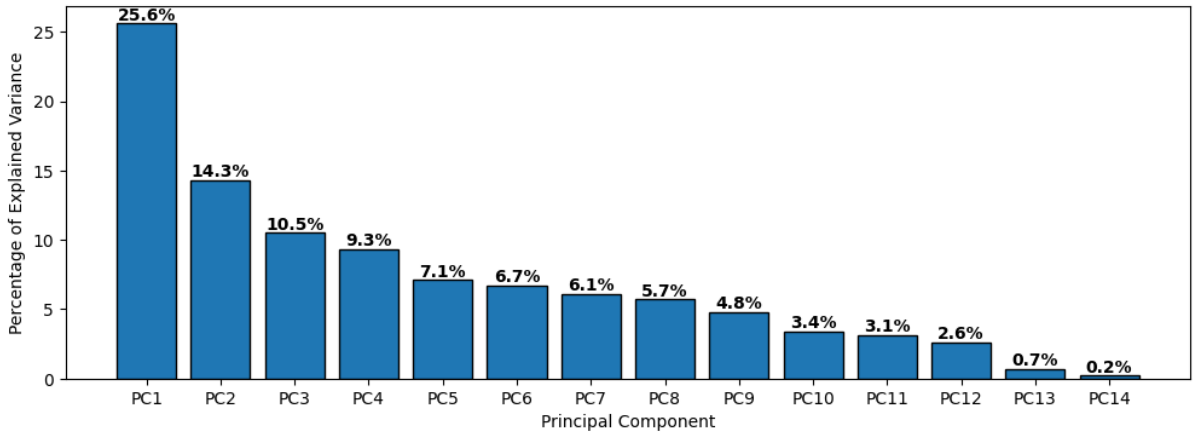
1. PCA Results

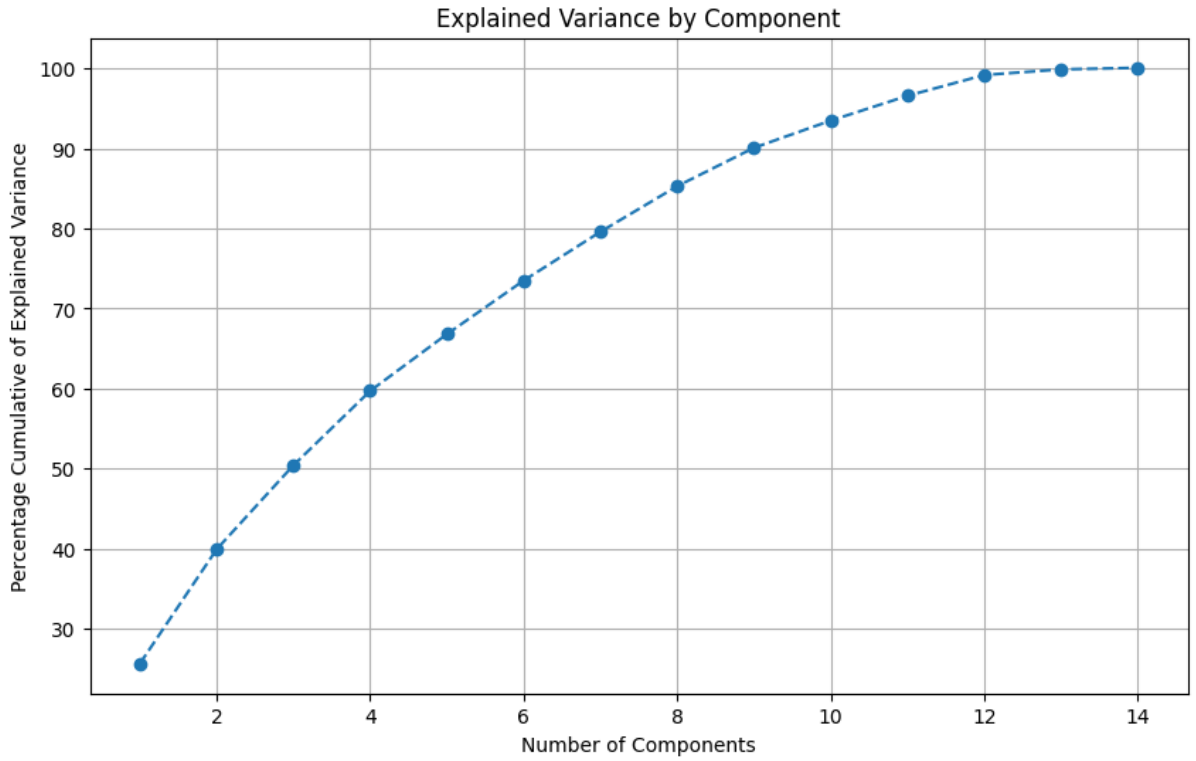
a. Without unit type





b. With unit type





2. CALCULATIONS

A. Without unit_type

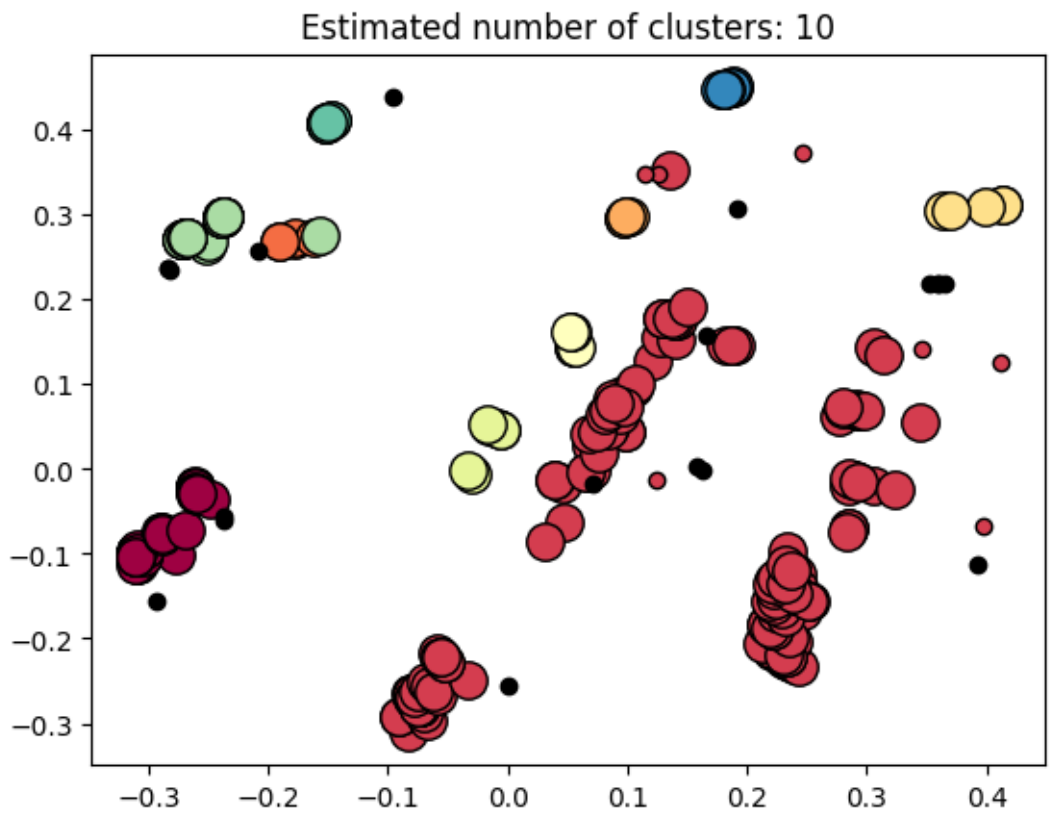
I. 7 Components ~85% explained variance

eps = 0.36, min_samples = 5

Clusters = 10

Outliers = 19

	0
PC1	[tv, building_staff, ac, wifi, breakfast]
PC2	[pool, building_staff, breakfast, area_name, tv]
PC3	[building_staff, pool, tv, area_name, wifi]
PC4	[breakfast, area_name, pool, ac, bathroom]
PC5	[area_name, breakfast, bedroom, bathroom, buil...
PC6	[parking, wifi, breakfast, ac, bathroom]
PC7	[ac, overall_rating, wifi, parking, tv]



segment

Cluster 2 177

Cluster 1 63

Cluster 8 29

Cluster 10 9

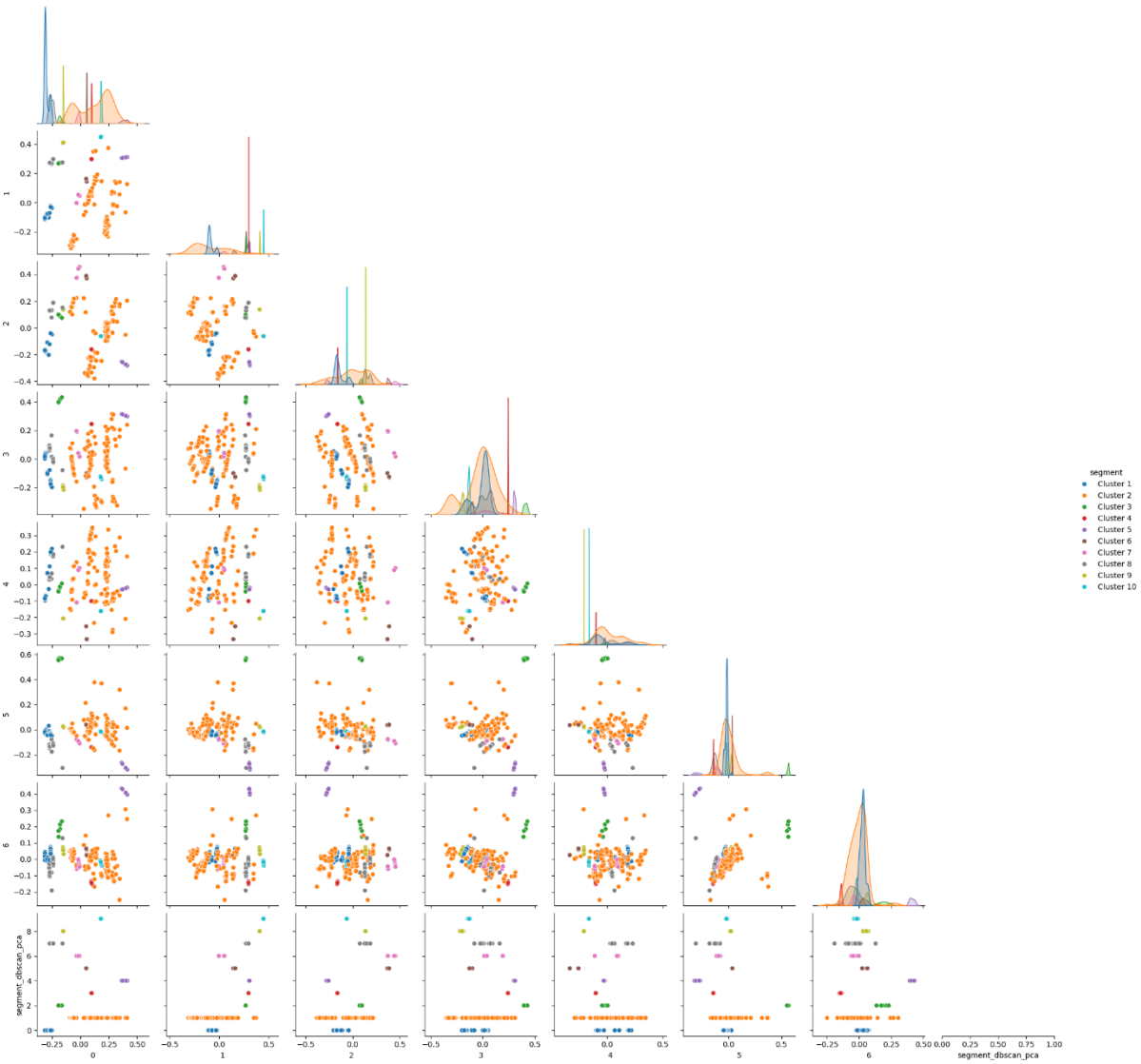
Cluster 7 7

Cluster 9 7

Cluster 3 5

Cluster 4 5

Cluster 5 5


segment
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Cluster 10
0
1
2
3
4
5
6
segment_dbscan_pca

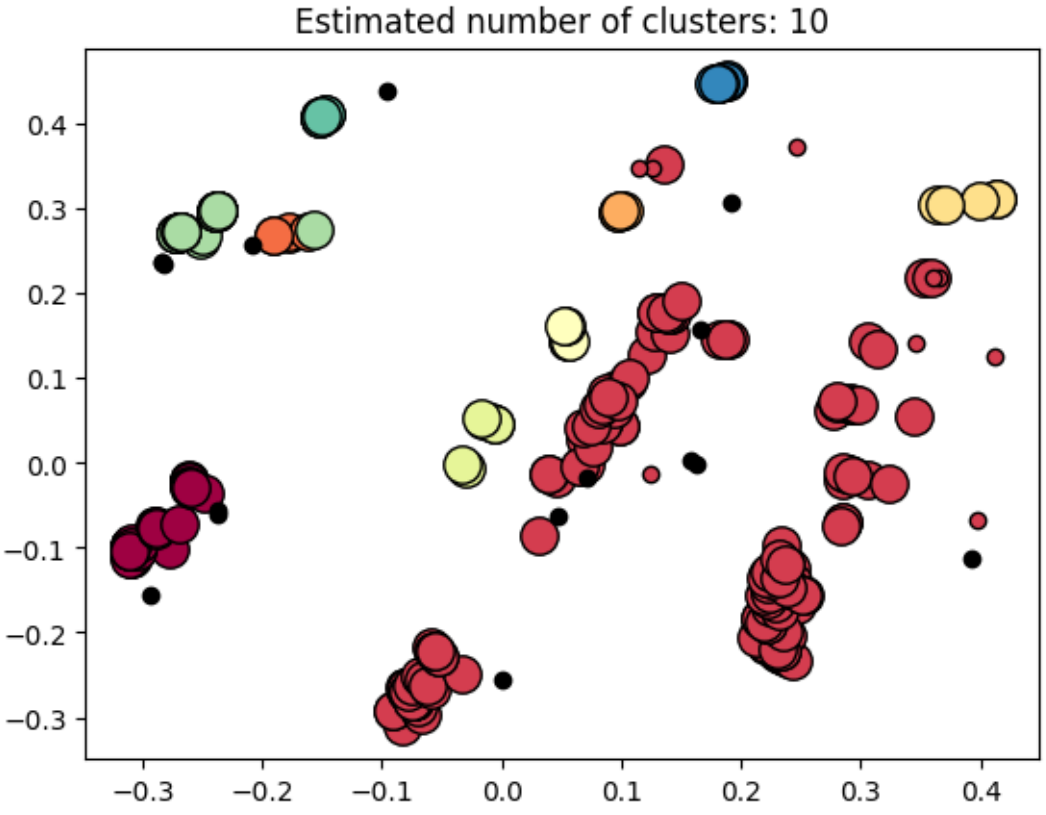
II. 8 Components ~90% explained variance

eps = 0.38, min_samples = 5

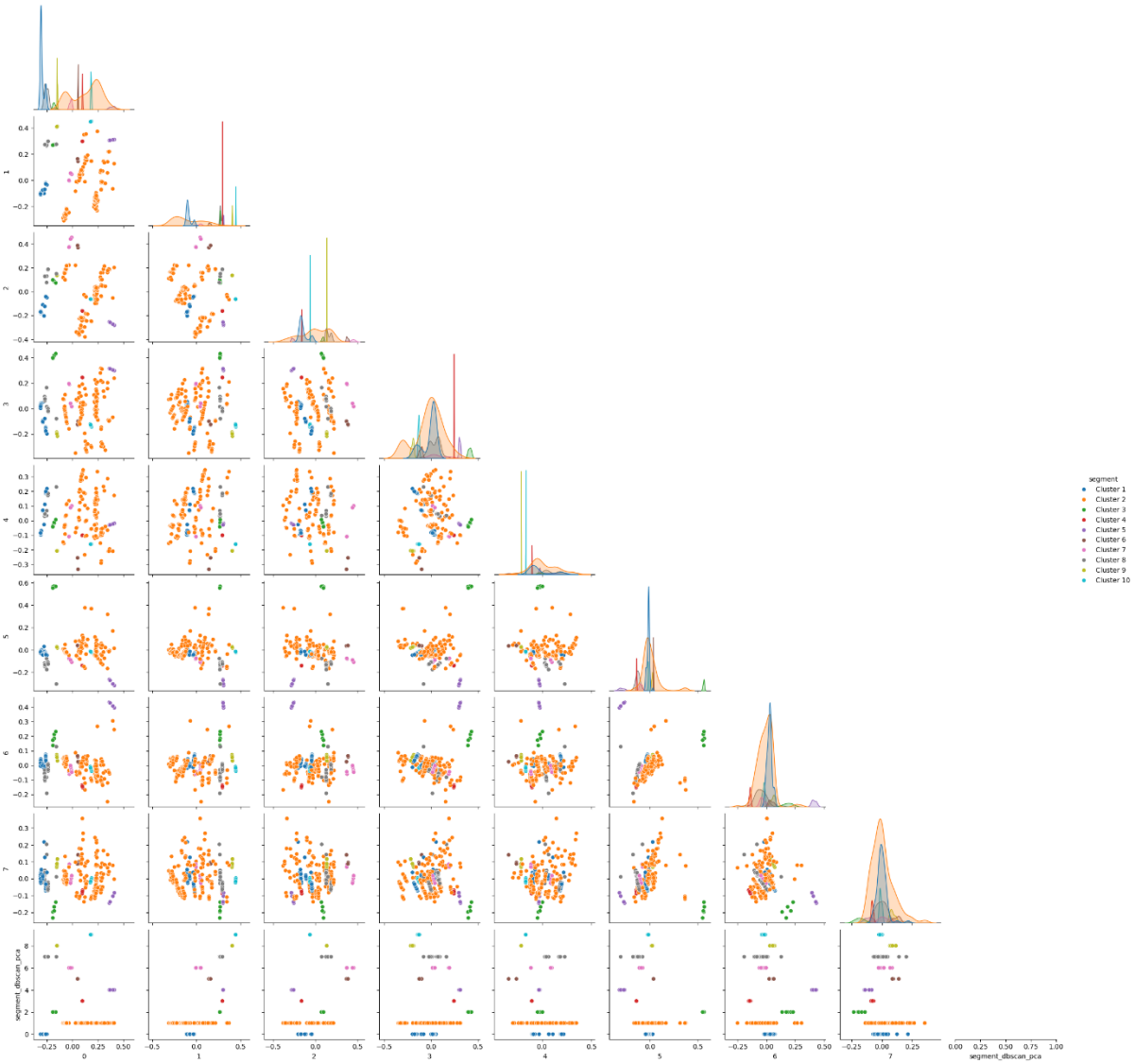
Clusters = 10

Outliers = 16

	0
PC1	[tv, building_staff, ac, wifi, breakfast]
PC2	[pool, building_staff, breakfast, area_name, tv]
PC3	[building_staff, pool, tv, area_name, wifi]
PC4	[breakfast, area_name, pool, ac, bathroom]
PC5	[area_name, breakfast, bedroom, bathroom, buil...
PC6	[parking, wifi, breakfast, ac, bathroom]
PC7	[ac, overall_rating, wifi, parking, tv]
PC8	[bedroom, beds, bathroom, overall_rating, park...



Cluster 2	180
Cluster 1	63
Cluster 8	29
Cluster 10	9
Cluster 7	7
Cluster 9	7
Cluster 3	5
Cluster 4	5
Cluster 5	5
Cluster 6	5



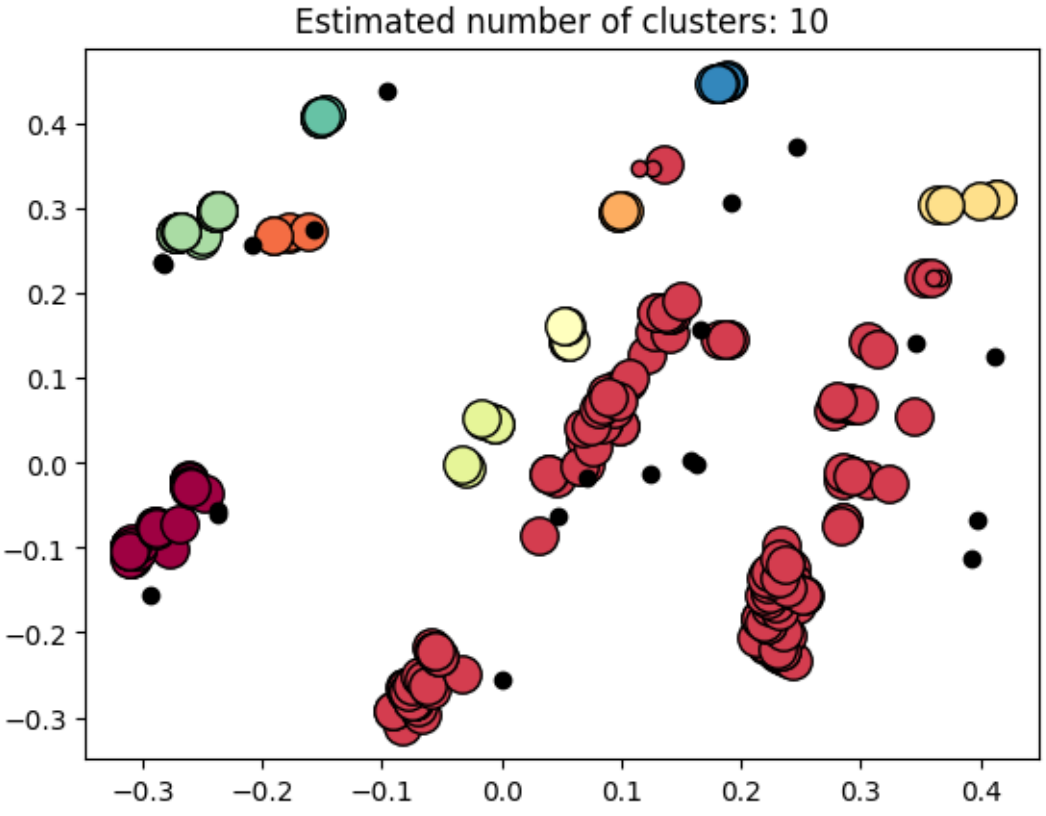
III. 9 Components ~95% explained variance

eps = 0.38, min_samples = 5

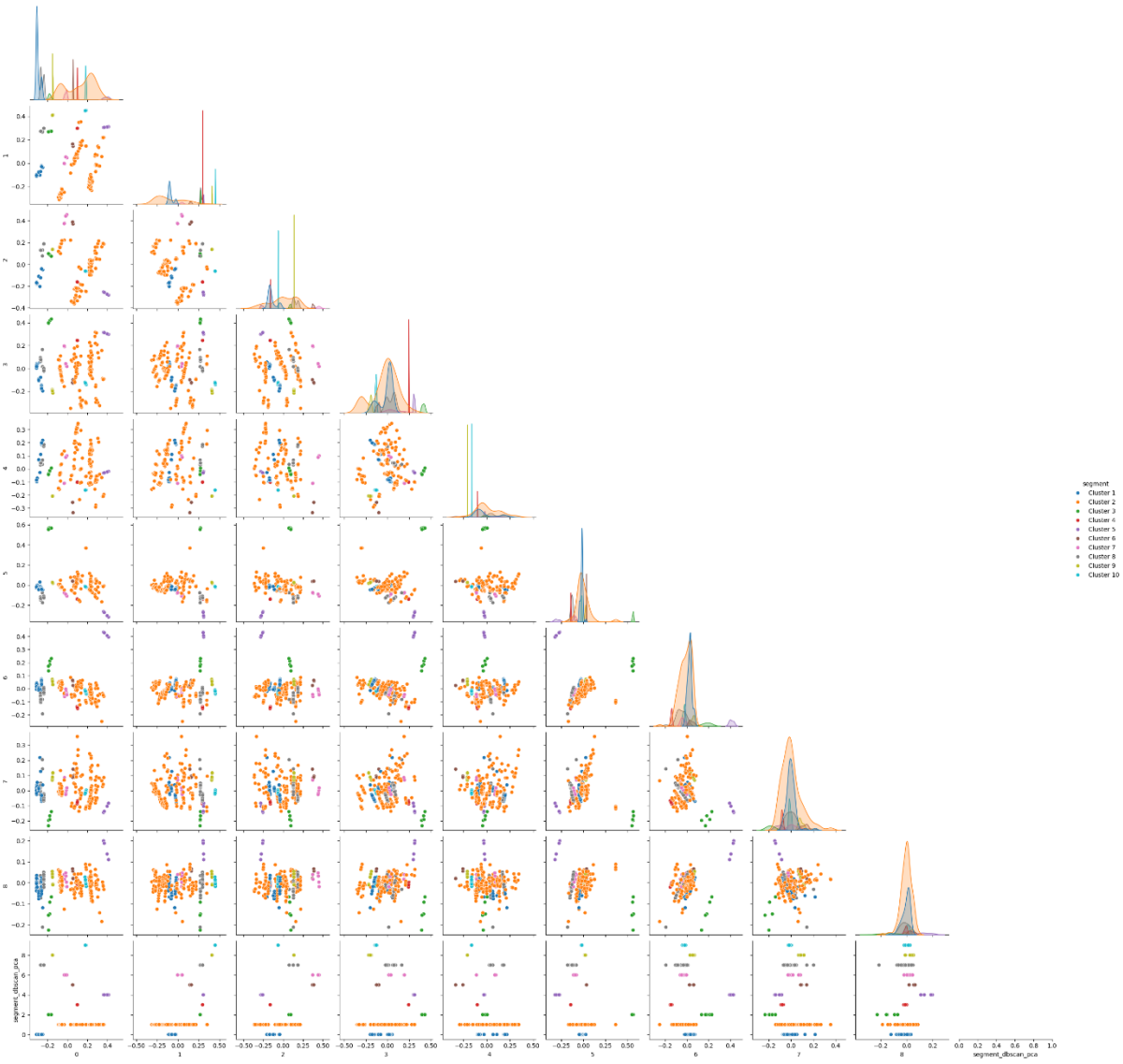
Clusters = 10

Outliers = 22

PC1	[tv, building_staff, ac, wifi, breakfast]
PC2	[pool, building_staff, breakfast, area_name, tv]
PC3	[building_staff, pool, tv, area_name, wifi]
PC4	[breakfast, area_name, pool, ac, bathroom]
PC5	[area_name, breakfast, bedroom, bathroom, buil...
PC6	[parking, wifi, breakfast, ac, bathroom]
PC7	[ac, overall_rating, wifi, parking, tv]
PC8	[bedroom, beds, bathroom, overall_rating, park...
PC9	[wifi, overall_rating, ac, parking, bedroom]



Cluster 2	175
Cluster 1	63
Cluster 8	28
Cluster 10	9
Cluster 7	7
Cluster 9	7
Cluster 3	5
Cluster 4	5
Cluster 5	5
Cluster 6	5



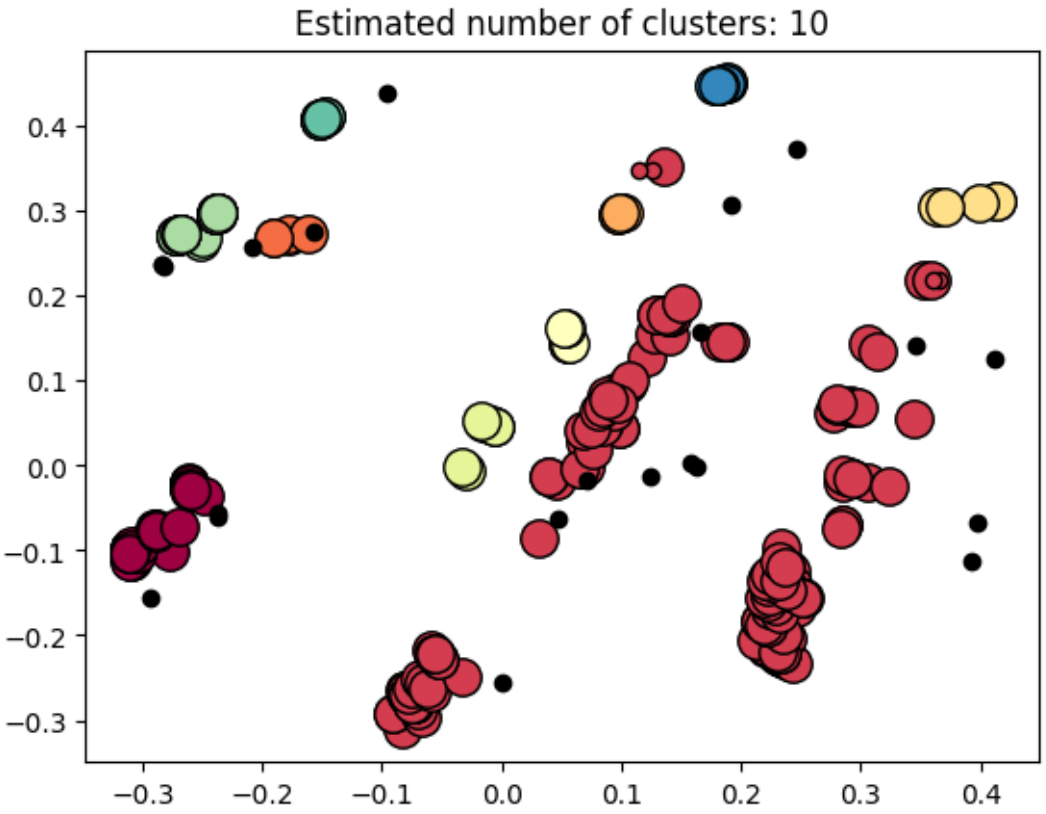
B. With unit_type

I. 7 Components ~80% explained variance

eps = 0.36, min_samples = 5

Clusters = 10

Outliers = 13



Cluster 2 175

Cluster 1 63

Cluster 8 28

Cluster 10 9

Cluster 7 7

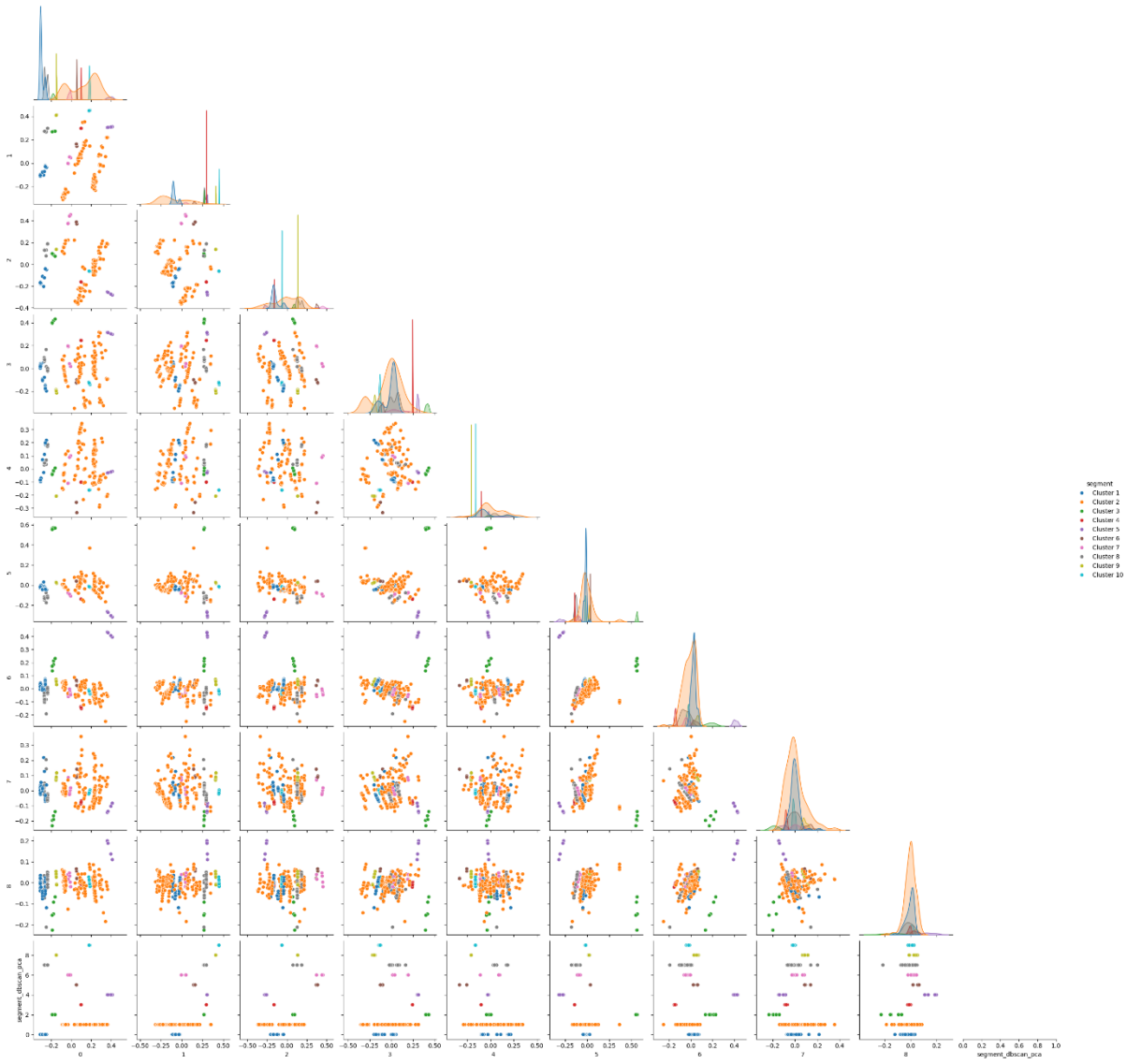
Cluster 9 7

Cluster 3 5

Cluster 4 5

Cluster 5 5

Cluster 6 5

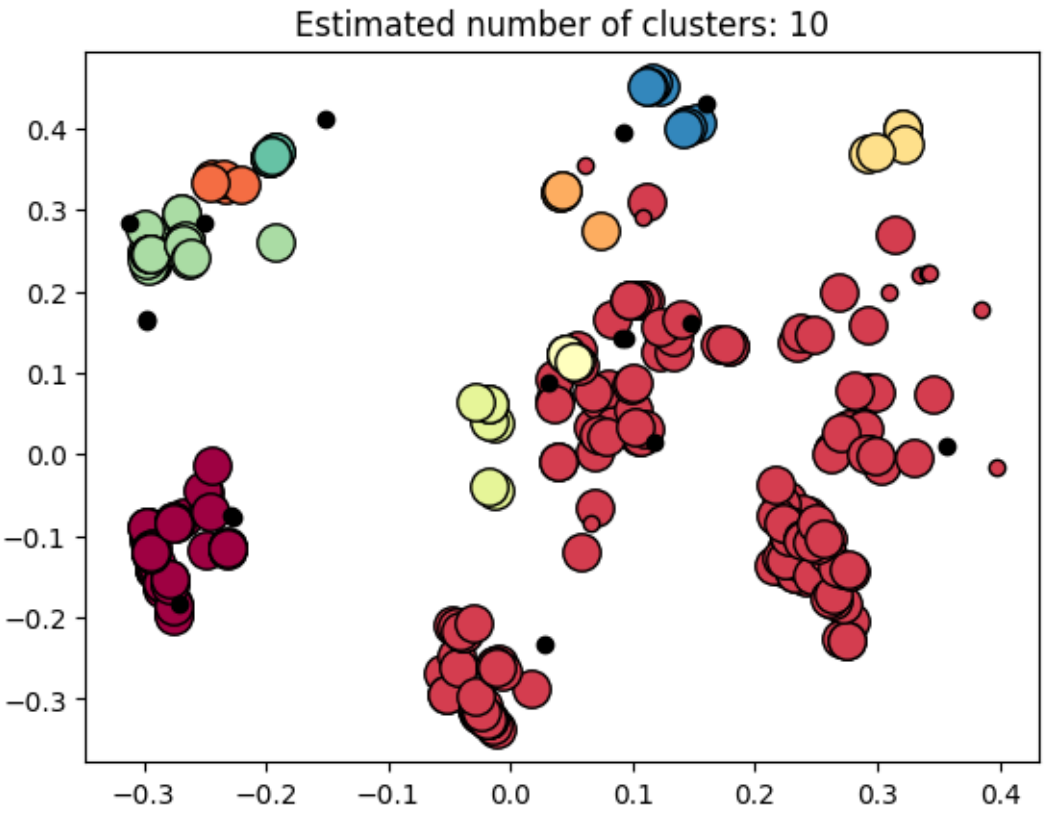


II. 8 Components ~85% explained variance

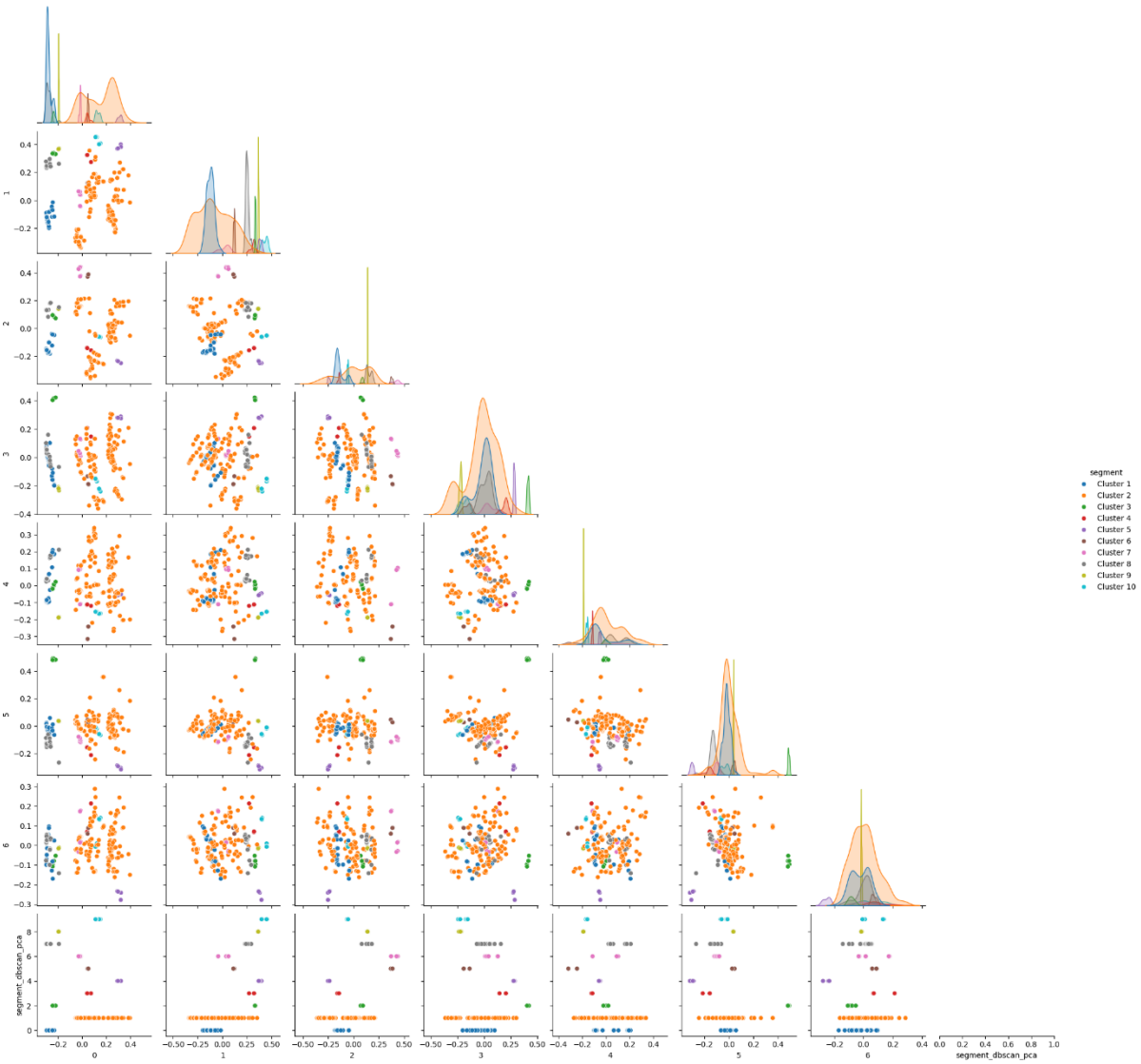
eps = 0.37, min_samples = 5

Clusters = 10

Outliers = 17



Cluster 2	179
Cluster 1	63
Cluster 8	29
Cluster 10	9
Cluster 7	7
Cluster 9	7
Cluster 3	5
Cluster 4	5
Cluster 5	5
Cluster 6	5

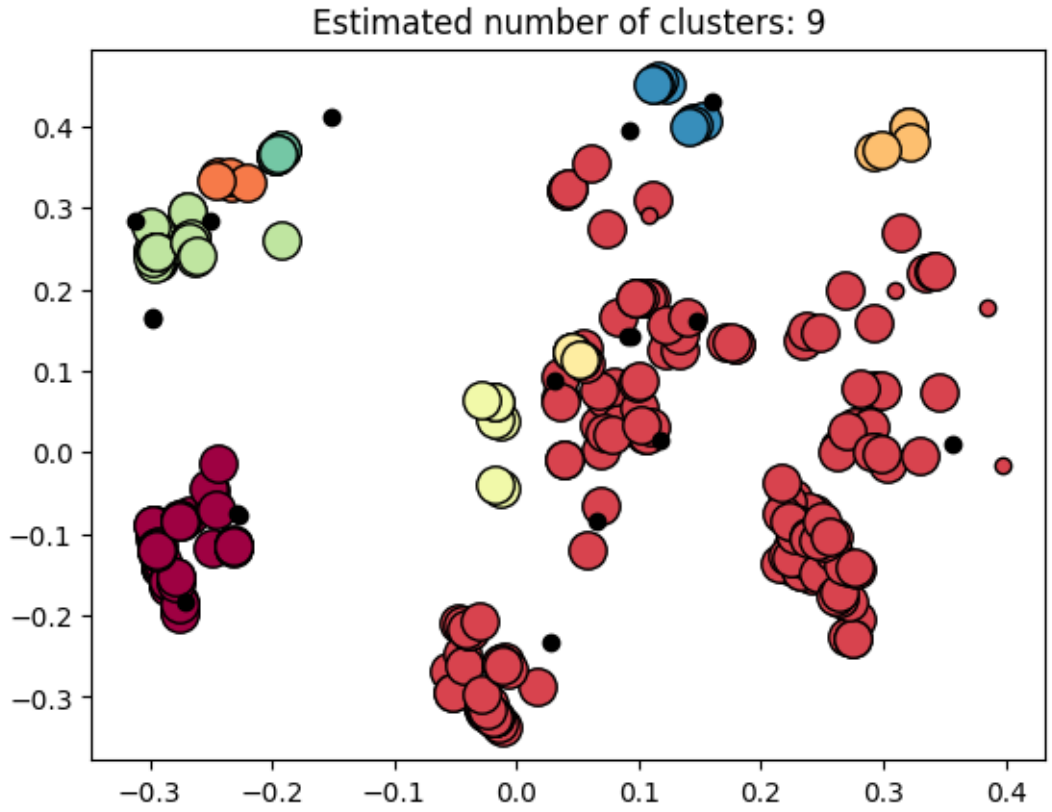

segment
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Cluster 10
0
1
2
3
4
5
6
segment_dbscan_pca

III. 9 Components ~90% explained variance

eps = 0.38, min_samples = 5

Clusters = 9

Outliers = 18



Cluster 2 183

Cluster 1 63

Cluster 7 29

Cluster 9 9

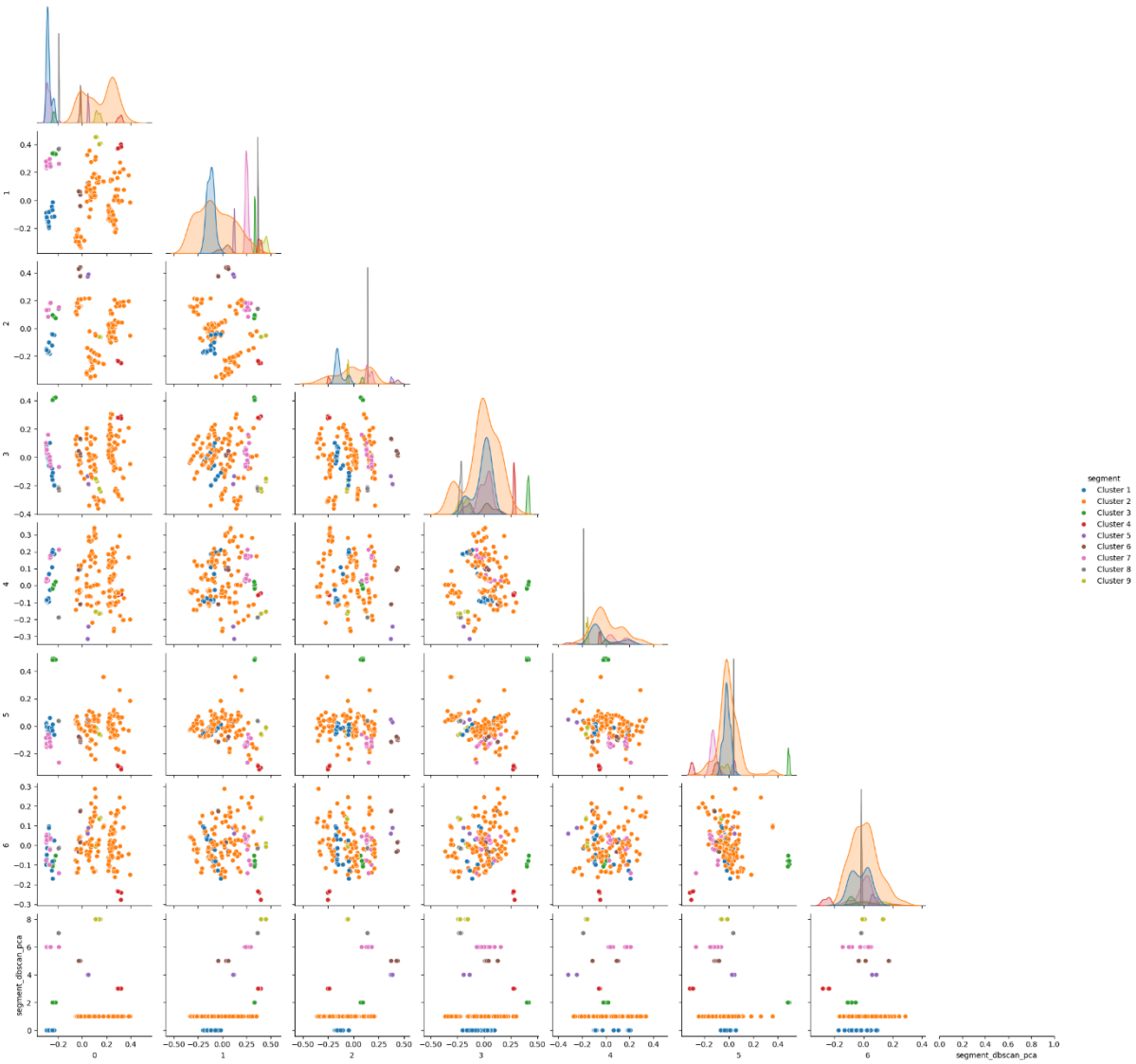
Cluster 6 7

Cluster 8 7

Cluster 3 5

Cluster 4 5

Cluster 5 5


segment
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
0
1
2
3
4
5
6
segment_dbscan_pca